Chapter 1: General introduction

Chapter 1

Cephalosporins

The discovery of cephalosporin C

In the summer of 1943, Giuseppe Brotzu, teacher of Hygiene at the University of Cagliari and Superintendent of Public Health during the campaign to eradicate malaria in Sardinia was watching the sea. He observed that the seawater near a sewer outlet was strongly polluted, but cleared up after 300m [1]. Although he attributed this to the purifying properties of the water, and was reluctant to step into the polluted water, his curiosity was too strong. So, one day he and his assistants stepped into the sewer to take samples. When grown on agar plates, the samples were shown to contain large quantities of *Salmonella typhimurium*, a pathogen that causes typhoid fever. This was not surprising, since the disease was endemic in Sardinia at the time. The reason for the pollution was evident. However, the question remained: Why does the water clear up? He continued with his research, and soon found that if the samples were grown on agar-glucose-peptone, the plates would be full of fungi within a few hours. He identified the fungus as *Acremonium cephalosporium*. Professor Brotzu, who was aware of the recent advances by Florey and Chain [2,3] on the penicillin research by Fleming [4], was convinced that the fungus produced some form of antibacterial agent. Indeed, the fungus was shown to inhibit the growth of *Salmonella*, and alcoholic extracts showed antibacterial effects against a series of pathogens. Subsequently, he successfully treated guinea pigs suffering from lesions, and some human patients soon thereafter. Unfortunately, he was unable to get support from the local government and pharmaceutical companies, even after he described his findings in the bulletin of the Institute of Health of Cagliari [5]. Therefore, he sent a sample of the fungus to Dr. Florey. In the London lab of Dr. Florey, three cephalosporins were purified from the fungus, named P, N and C. Of these, cephalosporin C (CPC, Figure 1) was shown to possess the greatest antibacterial activity [6].

The mechanism of action of β-lactam antibiotics

The cephalosporins belong to the family of β-lactam antibiotics. These are named after the reactive moiety of the compounds, the β-lactam ring. In CPC, the four-membered β-lactam ring is coupled to a six-membered dihydrothiazine ring to form the nucleus, 7-aminocephalosporanic acid (7-ACA), and a side chain, α-aminoadipic acid, is coupled via an amide bond to the nucleus (Figure 1).

β-lactam antibiotics target the synthesis of the peptidoglycan layer of the bacterial cell wall. Since peptidoglycan is only present in prokaryotic cells, toxicity to the eukaryotic cells of the patient is low. Peptidoglycan is a major component of the bacterial cell wall and is essential for the stability of the cell wall of both Gram-positive and Gram-negative bacteria. It consists of a disaccharide backbone formed by alternating β-1,4-*N*-acetylglucosamines and *N*-acetylmuramic acids. Peptides are linked to the latter component. These peptides are cross-linked for extra strength of the cell wall. As an example, in the cell wall of *Staphylococcus aureus*, the average chain length is in the range of 10 disaccharides. Pentapeptides consisting of L-alanine, D-glutamine, L-lysine and two D-alanines are attached to the *N*-acetylmuramic acids. The pentapeptides of two separate chains are cross-linked by the enzymatic insertion of a pentaglycine group.

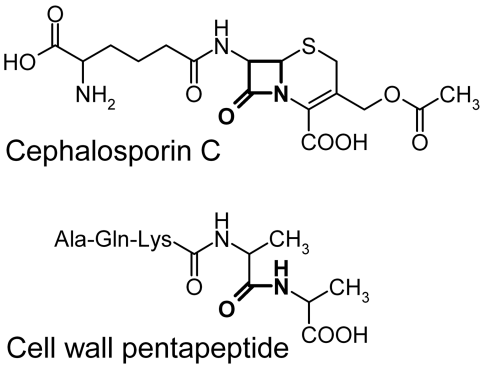

Figure 1: Structure of cephalosporin C and the cell wall pentapeptide.
The β-lactam ring of CPC is shown in bold. The structure closely resembles the two D-alanine residues of the pentapeptide that is cross-linked in the bacterial cell wall. The bacterial transpeptidase enzyme cuts the amide bond between the two alanines (shown in bold), thereby removing one D-alanine. Subsequently, it connects the rest of the pentapeptide to the lysine of a second pentapeptide via a pentaglycine. If the amide bond in the β-lactam ring of CPC is hydrolyzed, the dihydrothiazine ring remains attached and inhibits cross-linking and subsequent release of the enzyme.

Hereby, the lysine of one chain is connected to the first alanine of another chain. During this process the second alanine is removed from the pentapeptide (Figure 1).
The transpeptidase enzymes that perform the cross-linking also accept β-lactam antibiotics as substrates, because of the similarity of the antibiotic to the two D-alanines. The enzyme will break the β-lactam ring and during the reaction the antibiotic will bind covalently and irreversibly to the enzyme (Figure 1), thereby interfering with the proper construction of the bacterial cell wall, which will eventually lead to cell death [7,8].

The cephalosporins

CPC was shown to be active against a large number of pathogenic bacteria. However, its clinical use is hampered by a very low chemical stability. Consequently, cephalosporin derivatives were designed, which are more stable. Moreover, specific alterations to the cephalosporin compounds change activity profiles, increase activity towards resistant bacteria and increase resistance to stomach acid, which allows oral administration. Semi-synthetic cephalosporins are classified based on their activity profile, the antibacterial spectrum (Table 1, Figure 2) [9-11]. The first generation cephalosporins are active against Gram-positive microorganisms, such as staphylococci and streptococci. They are inactive against Gram-negative microorganisms, due to degradation by the β-lactamases of these bacteria. The antibacterial spectrum of the second generation cephalosporins is expanded with *Haemophilus influenzae* and some Gram-negative microorganisms, whereas the third generation is active against all Gram-negatives, buts lacks activity against some Gram-positive bacteria, mostly streptococci. Some classifications also include a fourth generation of cephalosporins that are active against virtually all Gram-negative and -positive bacteria.

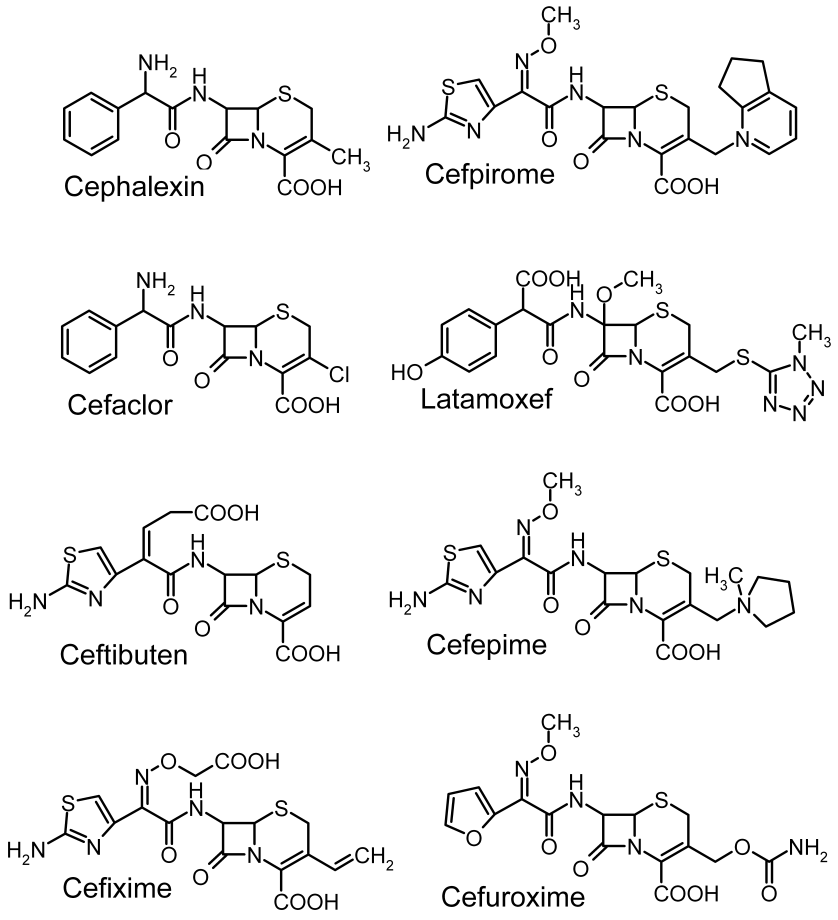

Figure 2: Structures of some semi-synthetic cephalosporins.

The total world market for cephalosporins was estimated to be approximately 10 billion US$ in 2000, and β-lactam antibiotics in general accounting for over 65% of the world antibiotic market [11]. According to IMS Health, cephalosporins as single preparation and in combination preparations are ranked 10 in the global drug sales in 2003 by an estimated sales of 8.3 billion US$, the highest ranking for any of the anti-infectives classes [12]. In The Netherlands, which usually has a restrictive drug use policy, antibiotic use in general is low, and cephalosporin use is very low. Semi-synthetic cephalosporins are regarded as "back up" drugs and are only to be administered when other drugs have failed, and after determination of the susceptibility pattern [9,10]. In contrast, cephalosporins are first choice antibacterials in many other countries. The widespread use has lead to the selection and propagation of resistant bacteria [13], but nevertheless, global cephalosporin use is huge.

Table 1: Marketed cephalosporins.
Oral: suited for oral administration. NL: Marketed in The Netherlands.
Data from the "Farmacotherapeutisch Kompas 2004" [9], "Informatorium Medicamentorum 2003" [10] and Elander [11]. *: Intriguingly, the classification used by Elander is different form that used by the other two sources. In case of discrepancies the classification of the first two sources is used. **: also classified as a fourth generation cephalosporin.

Generation	Name	Trivial name	Oral	NL
1	Cephalothin	Keflin		Yes
	Cephradine	Maxisporin, Velosef	Yes	Yes
	Cefadroxyl	Duracef, Moxacef	Yes	
	Cefazolin	Cefacidal, Kefzol		Yes
	Cephalexin	Keforal, Ceporex	Yes	Yes
	Cefaclor*	Ceclor	Yes	Yes
2	Cefuroxime	Cefofix, Zinnaf		Yes
	Cefuroxime axetil*	Zinnat		Yes
	Cefmetazole	Zefazone		Yes
	Cefonicid	Monocid		
	Cefoxitine	Mefoxin		Yes
	Cefamandol*	Mandol		Yes
3	Cefixime	Fixim	Yes	Yes
	Cefotetam*	Apatef		Yes
	Ceftibuten	Cedax	Yes	Yes
	Ceftriaxone	Rocephin		Yes
	Cefoperazone	Cefobine		Yes
	Cefotaxime	Claforan		Yes
	Cefprozil	Cefzil	Yes	
	Ceftazidime	Fortum, Kefadim		Yes
	Cefpodoxim	Orelox, Otreon	Yes	Yes
	Cefdinir	Omnicef	Yes	Yes
	Cefodizim	Modivid, Tinecef		Yes
	Ceftizoxim	Cefizox		Yes
	Latamoxef	Moxolactam, Moxam		Yes
	Cefepime**	Maxipime		
	Cefpirome**	Cefrom		Yes

Industrial production of semi-synthetic cephalosporins

The large quantities of semi-synthetic cephalosporins that are administered world wide have lead to the formation of a very competitive market that requires very efficient and cheap production routes. As is clear from the structures in Figure 2, all semi-synthetic cephalosporins can be produced from the cephalosporin nucleus, either 7-ACA or 7-aminodesacetoxycephalosporanic acid (7-ADCA) (see Figure 3). It was estimated that in 2000 approximately 1.95 million kg of 7-ADCA and 2.14 million kg of 7-ACA were produced [11].
Several production routes to the intermediates are depicted in Figure 3 [11,14,15]. The first process comprises the chemical deacylation of CPC. First the amino and carboxyl groups are protected, and then an iminochloride derivative is formed by a reaction with potassium pentachloride in the presence of base at -40ºC. Addition of alcohol results in

the iminoether, which is hydrolyzed to 7-ACA. A second chemical process involves the chemical modification of penicillin G (= phenylacetyl-6-aminopenicillanic acid or phenylacetyl-6-APA), produced by *Penicillium chrysogenum*, to 7-ADCA. This comprises sulfoxation of penicillin G using acetic acid followed by esterification, and expansion of the five-membered thiazolidine ring using pyridine, *N*,*N*'-bis(trimethylsilyl)urea and toluene. This results in phenylacetyl-7-ADCA, of which the side chain is removed by a chemical deacylation, similar to that used for CPC. Clearly, chemical production of the intermediates generates large quantities of waste and requires expensive and hazardous chemicals and reaction conditions.
In order to reduce costs and the environmental burden, research focused on enzymes to perform the required reactions. As soon as advances in enzyme production and recovery allowed, penicillin G acylase was used to perform the deacylation of phenylacetyl-7-ADCA to 7-ADCA. This enzyme removes the phenyl acetic acid side chain without modifying the β-lactam nucleus. Secondly, CPC is now converted to 7-ACA in a two-step enzymatic process. First the side chain is deaminated by a D-amino acid oxidase, resulting in an α-keto acid that spontaneously loses carbon dioxide in the presence of hydrogen peroxide to form glutaryl-7-ACA. Subsequent enzymatic deacylation of the glutaryl side chain yields 7-ACA. The enzyme used, cephalosporin acylase, removes a charged aliphatic side chain without damaging the β-lactam nucleus. Both these enzymatic processes have the advantage of generating less waste and requiring less expensive chemicals. However, the first enzymatic process is not yet totally "green" since the ring expansion is still a chemical reaction, and the latter enzymatic process is still relatively expensive.

New production routes for 7-ACA and 7-ADCA

Because of the drawbacks of the existing enzymatic processes, novel strategies are being investigated. One of these comprises the search for a cephalosporin acylase with activity towards the α-aminoadipyl side chain of CPC. Such a cephalosporin C acylase would allow a one-step enzymatic deacylation of this fermentation product. A second approach is to perform the ring expansion of penicillin G inside the *P. chrysogenum* fungus, allowing a fermentative production of phenylacetyl-7-ADCA, which can be deacylated in a one-step reaction by penicillin acylase. An enzyme for the expansion reaction was found in *Streptomyces clavuligerus*, in which it performs this reaction as part of the cephalosporin production pathway. Unfortunately, the enzyme does not accept penicillin G as substrate [14]. This problem was circumvented by feeding adipic acid to the fungus, resulting in the production of adipyl-6-APA, which is expanded to adipyl-7-ADCA [16]. However, this still left the search for an efficient adipyl acylase, since the existing cephalosporin acylases are not as efficient in the deacylation of this product as they are for the deacylation of glutaryl-7-ACA. Finally, a third strategy is the search for an expandase enzyme with a broad substrate specificity, so that the conversion of penicillin G to phenylacetyl-7-ADCA can be carried out enzymatically [17], resulting in a complete "green" production route when combined with a penicillin acylase. Eventually, equipping a *P. chrysogenum* with such an expandase gene may result in the fermentative production of phenylacetyl-7-ADCA.

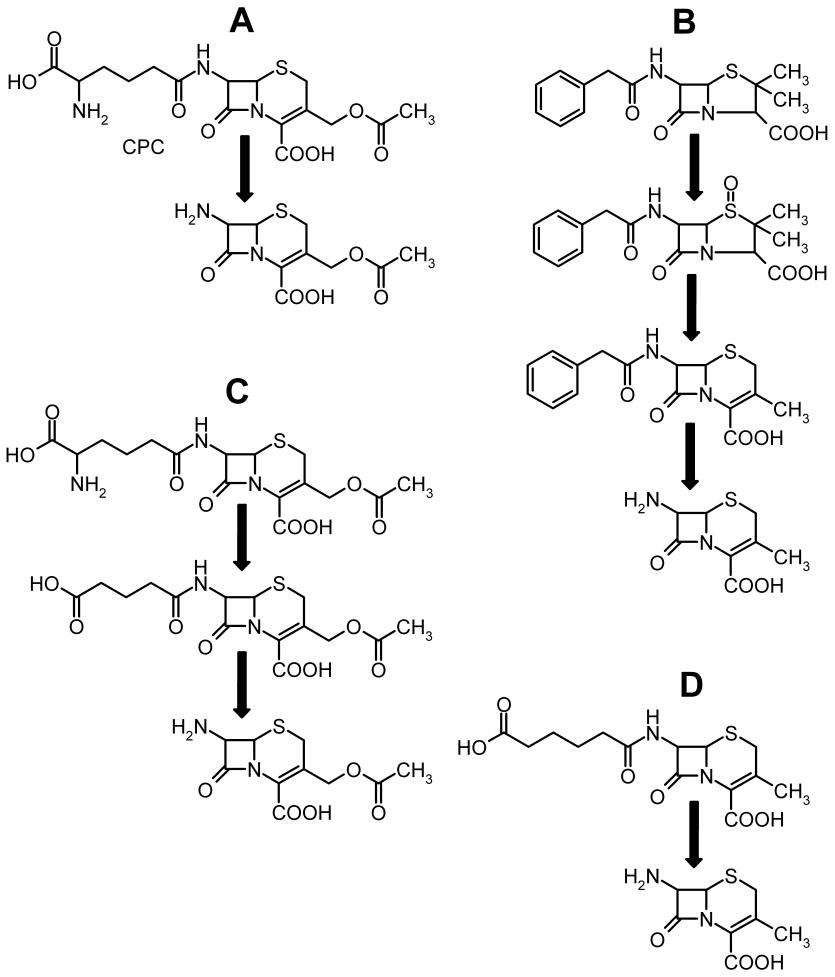

Figure 3: Industrial processes to 7-ACA and 7-ADCA.
Initially, production of these intermediates was solely accomplished by chemical means. Chemical deacylation of CPC yields 7-ACA (A), whereas sulfoxation of penicillin G, followed by a chemical ring expansion and deacylation of the phenylacetyl side chain yields 7-ADCA (B). Subsequently, chemical steps in the process were substituted by enzymatic reactions. Enzymatic deamination/decarboxylation of CPC results in glutaryl-7-ACA, of which the side chain is chemically removed to yield 7-ACA (C). Penicillin acylase could be employed to deacylate phenylacetyl-7-ADCA (the last step in B) and glutaryl-7-ACA could be hydrolyzed by cephalosporin acylases (the last step in C). A recent development is the enzymatic deacylation of adipyl-7-ADCA to yield 7-ADCA (D), and a enzymatic one-step deacylation of CPC to 7-ACA is currently under investigation (similar to A).

Chapter 1

Cephalosporin acylases

The cephalosporin acylases

The cephalosporin acylases are β-lactam acylases, highly specialized peptidases that are capable of cleaving the amide bond between a β-lactam nucleus and a side chain without damaging the β-lactam ring. β-lactam acylases are divided into two groups: those that accept substrates with aromatic side chains, also known as the penicillin acylases (EC 3.5.1.11), and those that accept substrates with a charged, aliphatic side chain, also known as the cephalosporin acylases (no separate EC entry, usually classified as well as 3.5.1.11). However, these trivial names can be confusing, as a cephalosporin with an aromatic side chain, e.g. phenylacetyl-7-ADCA, will not be hydrolyzed by a cephalosporin acylase, but is a substrate for the penicillin acylases. A proposal for a new classification based on side chain substrate specificity, e.g. dividing the β-lactam acylases into aromatic acylases and dicarboxylic acid acylases [18] has not been adopted by the other groups working in this area.

The physiological role of the β-lactam acylases is unknown. Deacylation of β-lactam antibiotics does not destroy the active β-lactam ring, so this cannot be their natural function. Penicillin acylases are presumably involved in nutrient scavenging [19], but there are no indications yet of the role of cephalosporin acylases. The β-lactam acylases are used commercially for the hydrolysis of β-lactams into the intermediates for the production of semi-synthetic penicillins and cephalosporins, 6-APA, 7-ACA and 7-ADCA (Figure 3). However, they can also be used for a variety of other reactions, including the addition of alternative side chains to β-lactam nuclei to form the clinically used semi-synthetic penicillins and cephalosporins, and the kinetic resolution of stereoisomers out of racemic mixtures by stereoselective acylation and/or deacylation. Recent key publications on alternative uses are the cephalosporin acylase catalyzed addition of the glutaryl side chain to amines at low water concentrations [20] and the use of cephalosporin acylase for resolution of stereoisomers [21]. The focus of this introduction will be on the use of cephalosporin acylases in hydrolysis, recent advances in other fields will be discussed in Chapter 2.

In 1981, Shibuya *et al* reported the results of a screening program for a glutaryl acylase. The penicillin acylases were quite well known by that time, and the first attempts at enzymatic hydrolysis of penicillin G at industrial scale had been performed, but no enzyme for the deacylation of CPC was available. However, the reports on the enzymatic deamination of CPC to glutaryl-7-ACA by D-amino acid oxidase prompted a screening program for an enzyme that deacylates this compound. First, the strains present in their laboratory were assayed for activity against glutaryl-7-ACA. Only one strain showed activity. The strain was identified as *Pseudomonas putida*, but activity was too low for commercial application. Subsequently, soil and wastewater samples were grown in medium with glutaryl-7-ACA added as the sole source of carbon. One strain, named SY-77-1, was selected with an appreciable activity. Analysis showed that this was a *Pseudomonas* strain, which was active towards cephalosporins with a glutaryl and a succinyl side chain. Activity towards substrates with an adipyl side chain was far less, and activity towards CPC was not detected. However, chemical treatment and selection resulted in the isolation of mutant GK-16, which produces the acylase constitutively in

Table 2: Cephalosporin acylases

Bacterium	Homologs	Activity succinyl	glutaryl	adipyl	CPC	Ref
SY-77	GK16, C427, sp130, KAC-1	++	+++	+	-	22-30
SE-83 *acy*I	V22-II	+	+++	+/-	-	31-33
SE-83 *acy*II	N176, V22	+	+++	++	+/-	31,32,34,35
B. laterosporus J1		+	+++	+++	-	34,35
B. laterosporus J1-II		+++	+	+		34
A14		+++	+	-	-	34,36
BL072			+++		-	37,38
Arthrobacter 45-8A			+++		+	39
P. syringae NBMCC1304			+++		-	40
Bacillus sp.			+			41
Aeromonas ACY95			-		+	42
Pseudomonas strain NCIMB40474			+			43
P. nitroreducens		-	+++		-	44
P. cepacia BY21			+++			45

high quantities but does not produce β-lactamase [22,23]. The acylase enzyme was purified from cell free extracts of GK-16 fermentation broths, showing a single band on polyacrylamide disc gel electrophoresis. The activity of intact cells and the efficient release of the enzyme by osmotic shock treatment and detergents suggested the enzyme to be located in the periplasmic space [24]. The gene encoding the acylase was isolated by screening for glutaryl acylase activity of a shotgun genomic DNA library of *Pseudomonas* GK-16. The enzyme was purified and showed two bands on SDS-PAGE of approximately 16 kDa and 54 kDa. DNA sequencing of the first 311 codons of the gene and N-terminal sequencing of the two subunits revealed that the gene encodes a precursor peptide that is transformed into two subunits. The small one was named α, the large one β. Furthermore, it was shown that the precursor starts with a stretch of 29 amino acids that are not present in the active enzyme. It was proposed that these constitute a signal sequence, which corroborated with the periplasmic localization of the enzyme [25].

After the description and analysis of the first cephalosporin acylase, many other followed. Another *Pseudomonas* with cephalosporin acylase activity was isolated from soil, strain SE-83. Two acylase genes were found in a shotgun genomic library, named SE-83 *acy*I and *acy*II. The second gene encodes an enzyme of two subunits of approximately 26 kDa and 57 kDa that deacylated adipyl-7-ACA quite well, and even showed some activity towards CPC [31]. Sequence analysis revealed that the *acy*II gene codes for a precursor containing the two subunits, but without a signal sequence, showing significant amino acid similarity with the GK-16 cephalosporin acylase. Differently, the *acy*I gene encodes a two subunit precursor that was dissimilar to the other two cephalosporin acylases. In contrast to these enzymes and penicillin acylase, the large subunit preceded the small subunit, and none of the two subunits started with the characteristic N-terminal serine. The AcyI enzyme shows predominantly activity towards glutaryl-7-ACA [32].

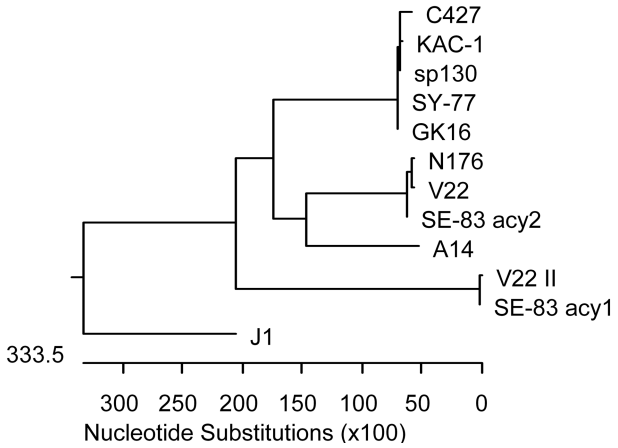

Figure 4: Phylogenetic tree of some cephalosporin acylases.
Protein sequences were aligned using the ClustalV method in MegAlign 5.07 (DNAStar, USA). Acylases are indicated by the strain number of the microorganism, for further data see the text. Some sequences can be found in GenBank, accession numbers AAN39264, AAA88424, P15558, BAA01048, S27199, AAF64242 and AAC34685. Distance between sequence pairs is indicated by the total length of the vertical connecting branches, while the scale below indicates the number of substitution events.

Subsequently, the cephalosporin acylases producing strains *Pseudomonas* A14, *Bacillus laterosporus* J1, *Pseudomonas* N176 and *Pseudomonas* V22 were isolated from soil [34]. The acylase gene from *B.laterosporus* J1 was shown to encode a highly dissimilar enzyme. The precursor of approximately 70 kDa consists of a single subunit prededed by a signal sequence, and the N-terminal residue is threonine instead of serine. The enzyme was equally active towards succinyl-7-ACA, glutaryl-7-ACA and adipyl-7-ACA, but did not show activity towards CPC [35]. The gene for *Pseudomonas* A14 cephalosporin acylase encoded for the usual propeptide containing two-subunits, an N-terminal serine and a signal sequence. The activity of this enzyme towards succinyl-7-ACA was much higher than the activity towards glutaryl-7-ACA, whereas no activity towards adipyl-7-ACA and CPC was detected. The cephalosporin acylases from *Pseudomonas* N176 and V22 were shown to be highly similar to the *Pseudomonas* SE83 AcyII enzyme, containing the same number of amino acids, less than 7% different amino acids, and comparable substrate specificities [36]. Soon, also a second cephalosporin acylase gene from *Pseudomonas* V22 was cloned and characterized, V22-II, which was virtually identical to *Pseudomonas* SE-83 AcyI [33].
Subsequent reports described a *Pseudomonas* SY-77 cephalosporin acylase homolog, the acylase from *Pseudomonas* C427 [26], the complete DNA sequence of the *Pseudomonas* GK-16 cephalosporin acylase [27], and another *Pseudomonas* SY-77 cephalosporin acylase homolog, the acylase from *Pseudomonas* sp.130 [28]. The latest cephalosporin acylase to be identified was that from *Pseudomonas* KAC-1 [29], which proved to be yet another homolog of the *Pseudomonas* SY-77 enzyme [30]. A phylogenetic tree of the

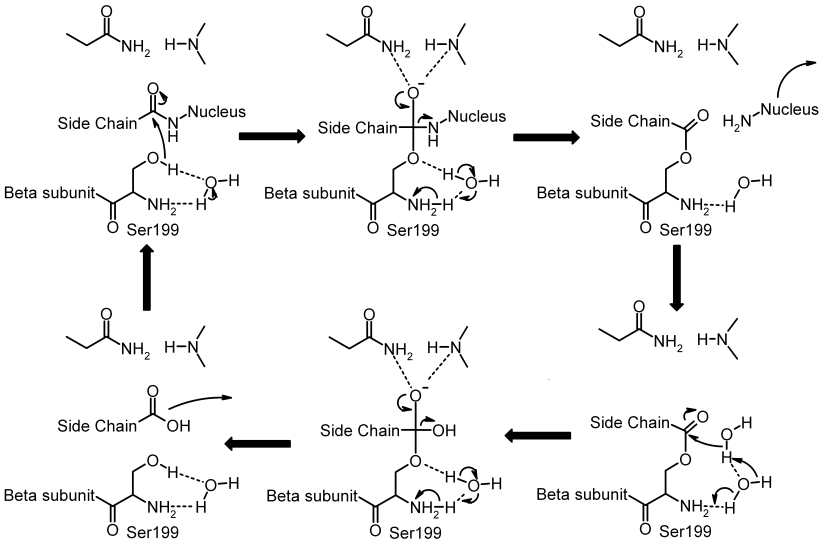

Figure 5: Proposed reaction mechanism of penicillin acylase, and probably cephalosporin acylase.
In acylase catalyzed hydrolysis the free amine group of the N-terminal serine plays a pivotal role. It is the base that initiates the reaction by activating the nucleophile, i.e. the hydroxyl group of the serine.

sequenced acylase genes is given in Figure 4. The SY-77 and SE-83 AcyII families are clearly distinguishable, but the other acylases are more less similar.
Many other cephalosporin acylase have been found, of which the gene has not been sequenced, including a second acylase from *B. laterosporus* J1 [34]. *Pseudomonas* strain BL072 was isolated from soil and showed glutaryl acylase activity, but a very low activity towards CPC [38]. The enzyme was purified and shown to consist of two subunits of approximately 24 kDa and 65 kDa [37]. Other acylases were found in *Arthrobacter* strain 45-8A [39], *Pseudomonas syringae* [40], *Aeromonas* ACY95 (which hydrolyzes CPC but not glutaryl-7-ACA)[42], an undetermined *Bacillus* species that produces an extracellular acylase [41], *Pseudomonas* strain NCIMB40474, which produces an acylase of two subunits of 18 kDa and 52 kDa [43], and *P. cepacia* BY21 [45]. A cephalosporin acylase was also isolated from cultures of *Pseudomonas nitroreducens*. The purified enzyme consists of two subunits of 35 kDa and 21 kDa, and the N-terminal sequence of the small subunit is somewhat similar to that of the *Pseudomonas* SE83 AcyI and *Pseudomonas* V22-II acylases. However, the N-terminal sequence of the large subunit is not similar to that of the other two acylases [44].
Finally, strains showing cephalosporin acylase activity, which have not been analyzed in more detail, are *Bacillus cereus*, *Achromobacter xylosooxidans*, *Bacillus* sp., *Pseudomonas* sp., *Pseudomonas paucimobilis* [46] and *Flavobacterium* sp.650 [47].
Clearly, cephalosporin acylases are to be found in many microorganisms. It should be noted that the acylases discussed in this introduction are all described in peer-reviewed journals. A great many more have been reported just in patents because of the economic

value of these enzymes, and have not been discussed here. Another point that should be noted is that most reports have not been followed up by later articles. Research has focused on the cephalosporin acylases of the SY-77 and the SE-83 AcyII families.

Biochemical properties of cephalosporin acylases

The cephalosporin acylases of the SY-77 family are studied most extensively. The gene product is a propeptide consisting of a signal sequence, the α-subunit, a spacer peptide, and the β-subunit. The signal sequence directs the propeptide to the periplasm of the native *Pseudomonas* species, and is thereafter removed. Subsequently, the enzyme folds and the amide bond between the spacer peptide and the β-subunit is cleaved by autocatalytic processing, in what is thought to be an intramolecular event. Finally, the amide bond between the α-subunit and the spacer peptide is cleaved in what is thought to be an intermolecular event. The spacer peptide is released, and the active enzyme consisting of the α- and β-subunit remains [48]. Ser199 becomes the N-terminal residue of the β-subunit (many different systems for amino acid numbering of β-lactam acylases exist, throughout this thesis numbering is consecutively starting from the first residue of the gene translation product, Met1 of the signal peptide). This residue is essential for autocatalytic processing and hydrolytic activity, which is consistent with the data from other members of the N-terminal nucleophile hydrolase superfamily [49,50]. The 3D structures of the cephalosporin acylase from *Pseudomonas* KAC-1 and *Pseudomonas* GK-16 showed the presence of an αββα sandwich motif at the active site of the enzyme, which is a second characteristic of the N-terminal hydrolase superfamily [51-53]. The proposed model for the mechanism of autocatalytic processing was later confirmed by analysis of the 3D structures of mutants that do not maturate correctly [53-55].
The *Pseudomonas* SE-83 family has also been studied in detail, mostly using mutagenesis to increase the CPC acylase activity (see below). A preliminary report on the crystallization and X-ray analysis of the N176 acylase was published in 2000 [56], but has not been followed by a report of the completed 3D structure. The other cephalosporin acylases have not been studied further, and this introduction will focus on the SY-77 acylase family.
It was assumed that the catalytic mechanism of cephalosporin acylase was similar to that of penicillin acylase [57] since the N-terminal serine of the β-subunit of cephalosporin acylase was also found to be essential for catalytic activity [48]. The proposed reaction mechanism of penicillin acylase is shown in Figure 5. At the active site of the enzyme, the uncoupled amine group of the N-terminal serine activates the hydroxyl group of the same residue via a bridging water molecule. The oxygen of the hydroxyl group then performs a nucleophilic attack on the carbon atom of the amide bond of the substrate. A tetrahedral intermediate is formed, which is stabilized by an oxyanion hole, consisting of the side chain amide of Asn442 and the main chain nitrogen of Val268 [51,52]. The intermediate collapses, resulting in a liberated β-lactam nucleus and a covalently bound acyl-enzyme. The carbon atom of the ester bond between side chain and enzyme is then attacked by a water molecule, during which a similar tertrahedral intermediate is formed. This collapses, resulting in the regenerated enzyme and a liberated side chain, the overall reaction being a hydrolysis of the amide bond between nucleus and side chain [57]. Recent results indicate that the rate-limiting step of the reaction is the formation of the acyl-enzyme. This means that the kinetic parameters are determined by the binding of the substrate to the enzyme

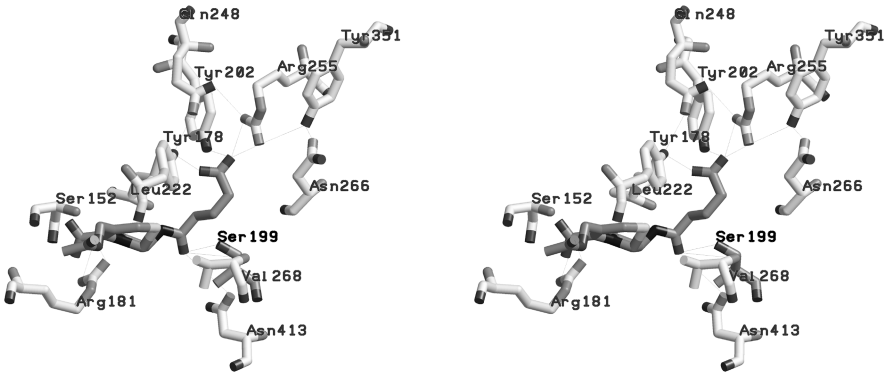

Figure 6: Selected residues of cephalosporin acylase that are directly or indirectly involved in substrate binding.
The substrate glutaryl-7-ACA and the catalytically active residue Ser199 are shown in dark grey. Hydrogen bonds are depicted as dashed lines. The carboxylate group of the side chain is part of an intricate network of hydrogen bonds. Color figure at end of thesis.

and the formation of the tetrahedral intermediate [58]. Recent 3D structure analysis shows that the active sites of the penicillin and cephalosporin acylases are highly similar [51,52], supporting the idea of similar reaction mechanisms. It seems therefore logical to assume that also for cephalosporin acylase the rate limiting step in the deacylation is the formation of the acyl-enzyme. This information is crucial for interpretation and prediction of the effect of mutations in the enzyme on hydrolysis of glutaryl-7-ACA and alternative substrates.

The structure of cephalosporin acylase

During the last two decades, great effort has been put in the determination of the 3D structure of the cephalosporin acylases of the SY-77 and SE-83 AcyII families. The determination of the 3D structure of *P. diminuta* KAC-1 cephalosporin acylase was reported in 2000 [51], followed by the structure of enzyme bound to glutaryl-7-ACA one year later [59]. At last, the interactions between substrate and enzyme were known, which would greatly facilitate the generation of cephalosporin acylases with altered substrate specificities. In Figure 6 a selection of the residues of cephalosporin acylase surrounding the substrate are shown. These are the residues that interact with the substrate, with Ser199 as the catalytically active N-terminal serine, the first residue of the β-subunit. The acetoxy group at the 3' position of the nucleus has hydrophobic interactions with Ser181 of the α-subunit, and forms hydrogen bonds to the guanidine group of Arg184. The double rings of the nucleus have hydrophobic interactions with the aromatic ring of Tyr178, Leu222 and Val268. Finally, the aliphatic part of the side chain has hydrophobic interactions with Leu222, Val268 and Phe375, whereas the carboxylate head of the side chain engages in a network of hydrogen bonds between Tyr178, Tyr231, Gln248, Arg255, Asn266 and Tyr351. Clearly, the interactions of the enzyme with the side chain are far

more numerous than the interactions with the nucleus, especially if not using 7-ACA as the nucleus but 7-ADCA, which lacks the acetoxy group.

Methods to find novel biocatalysts

"Industrial process would like to meet an enzyme"

In order to make the one-step enzymatic production of 7-ACA and 7-ADCA from fermentation products industrially viable a β-lactam acylase is required that is very efficient in hydrolyzing CPC or adipyl-7-ADCA. However, from the former section it is clear that all well-characterized cephalosporin acylases are highly active towards glutaryl-7-ACA, but less or far less active towards adipyl-7-ADCA, and far less or basically not active against CPC. Process design for the enzymatic deacylation of adipyl-7-ADCA has been described in detail [60-63] and a higher activity of the enzyme is desired.
In order to obtain a more effective biocatalyst, four strategies may be used:
1. Screening of samples from natural sources
2. Random mutagenesis of available enzymes
3. Rational mutagenesis of available enzymes
4. Screening for homologous genes (datamining of genome sequence projects)

Screening of samples from natural sources

Most currently known cephalosporin acylases have been discovered by large scale screening of samples obtained from soil. The most simple procedure is to isolate single colonies of microorganisms on agar plates, cultivate them and analyze their performance in the desired enzymatic reaction (e.g. [22,29]). This can amount to rather huge projects, in which 100,000 strains are assayed by HPLC [26]. The use of derivatives of the desired substrate that are easier to assay, such as glutaryl-*p*-nitroanilide and glutarylcoumarin instead of glutaryl-7-ACA, will result in a large number of false positives [38,46], but related β-lactam substrates such as glutaryl-7-ADCA and cephalothin have given good results [38].
It is more efficient to select microorganisms from soil samples than to screen isolated microorganisms. The most basic selection method is to grow the samples in medium containing the desired substrate as the sole source of nitrogen or carbon [64]. However, this selection is not very enzyme type specific, and should only be used as a first selection round. Acylases for glutaryl-7-ACA can be selected for by adding to the medium glutaryl-*p*-nitroanilide, which turns yellow upon deacylation [45]. Screening on plates containing glutaryl-6-APA or adipyl-6-APA that are overlaid with the 6-APA sensitive strain *Serratia marescens* has also been reported [47]. Unfortunately, derivatives of α-aminoadipyl, which could be used for screening for a CPC acylase, are very difficult to synthesize [65], and most are unstable, e.g. α-aminoadipyl-*p*-nitroanilide [66]. Microorganisms selected on the derivatives of *p*-aminobenzoic acid and 4-methoxy-β-naphthylamine showed a low activity towards CPC [66]. Clearly, the identification of efficient adipyl and CPC acylases using these methods will be difficult. With regard to selection on substrate derivatives, the relatively few interactions of *Pseudomonas* SY-77 cephalosporin acylase with the nucleus is a big advantage. The chance is relatively high

that activity towards the selection substrate and activity towards the β-lactam substrate are correlated.
In spite of much effort no efficient CPC acylase has been isolated from natural sources. Moreover, most microorganisms present in nature cannot be cultivated in the laboratory, and their enzymes are therefore not investigated. Recently, a new technique called metagenome sampling has been generated, in which the total DNA from a sample is isolated, cloned into expression vectors, and expressed in hosts such as *E. coli*. This creates libraries that can be screened or selected for the desired function [67]. It may be that in spite of the difficulties with expression of genes in heterologous hosts, novel enzymes, e.g. CPC acylases, will be found in this way.

Random mutagenesis of available enzymes

Protein function is determined by the amino acids of the protein, which are determined by the basepairs of the encoding gene(s). Creating diversity in the gene will result in proteins with different characteristics, for instance an altered substrate specificity of an enzyme. This can be used to create and improve an adipyl or CPC acylase. In order to limit the alterations required to generate an adipyl or CPC acylase it is logical to start from an enzyme that mimics the desired function as close as possible. The high similarity of the glutaryl, adipyl and α-aminoadipyl side chains (see Figure 3) suggest that the best starting point is one of the available cephalosporin acylases.
There are several methods to generate random DNA diversity. The most widely used way is the introduction of point mutations, for example by error-prone PCR or spiked oligo mutagenesis. These methods generate mutations in one or more basepairs of the gene. However, since amino acids are encoded by codons of three basepairs each and since it is highly unlikely that multiple point mutations will occur in one codon, an amino acid will only be changed into a limited set of other amino acids, not all 20. Mutagenesis of entire codons is needed for complete randomization. In saturation mutagenesis all basepairs of a codon are mutated simultaneously, creating all possible 20 amino acids at the position. However, the number of mutants will increase exponentially for each mutated codon. Clearly, there will always be a trade-off between the sequence space that is investigated, and the number of mutants that have to be analyzed. A third way is recombination, of which DNA shuffling is the best known example [68]. This comprises the recombination of variants of a single gene, or of homologous genes. The mutations or differences of multiple good enzymes are combined, which typically results in a pool of worse, equal, and sometimes better enzymes.
All of the abovementioned methods will result in mutant libraries, which will have to be screened or selected for the desired function. Preferably a specific host strain is chosen that allows simpler analysis. A method used in this thesis is growth selection. The gene is expressed in a bacterium that needs leucine from the medium to propagate, but is given adipyl-leucine instead. Therefore, colonies will only grow if the expressed gene is capable of removing the adipyl moiety from the leucine, i.e. if the expressed gene is an adipyl acylase. Of course, there are some restrictions to this method: a) an active enzyme must be formed in the host, b) the selection substrate must not be lethal to the host, c) the selection substrate must not be hydrolyzed by other enzymes of the host, d) the investigated enzyme must have access to the selection substrate, e) the liberated leucine must only be available

to the cell expressing the hydrolyzing enzyme, and f) the activity against the selection substrate must also imply activity towards the actual substrate, e.g. adipyl-7-ADCA.
Another selection method is phage display that offers the advantages of coupling phenotype and genotype combined with analysis of the protein. The mutant libraries are cloned in a phagemid, which will result in alibrary of phages presenting on its surface a protein, which is encoded by the variant of the gene that was cloned into the phagemid. This allows for the selection of proteins with the desired function together with the corresponding gene. However, selection of phage display libraries based on affinity is straightforward, but the selection of mutants with an increased catalytic activity is far more demanding [69,70]. Phage display of penicillin acylase has been achieved [71], but a successful selection of mutant acylases has not yet been reported.

Rational mutagenesis of available enzymes

Mutations that alter the substrate specificity of an enzyme can also be introduced intentionally. Rational mutagenesis usually requires both information about the 3D structure of the enzyme and preferably the enzyme-substrate complex, and some insight on the structure-function relationship. It is therefore very information-intensive. Rational design can be a very effective strategy to alter the enzyme mechanism, substrate specificity and stability [72]. However, even when a good 3D structure is available, not every attempt will be successful [73]. Some systematic problems exist. First, “irrational” mutations that can have a strong effect will be neglected [72]. Second, some parameters are usually approximated in order to decrease computer time, such as the use of force fields instead of quantum mechanics for atom and molecular properties, using constants instead of calculating the effect of solvent molecules, and scanning for energy minima instead of calculating the energy of each possible configuration. Clearly, the quality of the rational design will depend heavily on the quality of the used approximations, which can differ for each situation [74]. Nevertheless, some very good results have been achieved [72-74] and the computational power of publicly available computers is rapidly increasing, as is the quality of modeling software. The recent determination of the crystal structure of a SY-77 type cephalosporin acylase has greatly increased the possibilities for rational mutagenesis. Another strategy is de novo enzyme design, which comprises the generation of a biocatalyst from an inactive scaffold protein. Advances are made in this field [75], but the approach remains very challenging as the interactions needed for a complex enzymatic reaction can hardly be mimicked.

Screening of homologous genes

Improvements in DNA sequencing technology have enabled large-scale whole genome sequencing [76,77]. A number of the finished and unfinished sequences are publicly available at the NCBI website (http://www.ncbi.nlm.nih.gov/BLAST/). Via a so-called BLAST search [78] genes or proteins similar to a query gene or protein can be identified. Datamining of the available genome sequences will result in a number of candidate genes that might exhibit the same or related function as the target gene. Analysis of these genes or the microorganisms is basically a special way of screening of samples from natural sources. A restriction of this method is that a query gene is required, which must show a function that is related to the desired function. Searching for a protein with a completely

novel function is impossible. Another drawback is that a related enzyme may catalyze a related reaction on a very different substrate. Recently, an enzyme from *Ralstonia* XJ12B was shown to hydrolyze the amide bond between the side chain and nucleus of acyl-homoserine lactone molecules, which are signal molecules in a biological process called quorum sensing [79]. The enzyme is a cephalosporin acylase homolog, but does not hydrolyze β-lactam compounds.

The search for a biocatalyst for the one-step enzymatic hydrolysis of adipyl-7-ADCA and CPC

Although most of the available cephalosporin acylases have been found by screening of samples from natural sources, this has not yet resulted in a highly active adipyl or CPC acylase. Random mutagenesis of penicillin acylase by chemicals and UV light followed by growth selection using glutaryl-leucine has resulted in mutants that were highly active towards the selection substrate [80]. However, the mutation of the acylase increased the interaction with the leucine moiety of the selection substrate, so the growth on glutaryl-leucine was not because of a glutaryl acylase activity [81]. Not surprisingly, no activity of the mutant towards glutaryl-7-ACA and other cephalosporins has been reported.
Rational mutagenesis of a member of the SE83 cephalosporin acylase family has been tried to increase the activity towards CPC. Residues were picked based on chemical modification studies and mutated to a number of other residues, resulting in mutants showing a two-fold increase of CPC acylase activity, at best [82-85]. The lack of a 3D structure has hampered further rational mutagenesis of this enzyme.
In contrast, the 3D structure determination of a SY-77 cephalosporin acylase family member and the enzyme-substrate complex has greatly stimulated rational mutagenesis. The first report, on the structure of the apoenzyme, was already accompanied by suggestions for residues to be modified in order to create a CPC acylase [51]. The report on the structure of the enzyme-substrate complex was accompanied by suggestions towards which amino acids these residues should be changed [59]. The structure of the acylase with a model of the substrate was accompanied by a list of specific mutations that would generate a CPC acylase [52]. Disappointingly, mutants with an increased activity towards CPC based on these suggestions have not been published yet. The single exception is formed by the report of Oh *et al* on random mutagenesis of residues that were rationally selected based on substrate modeling. The best mutant showed an eight-fold increase in CPC hydrolysis, which is still more than 10-fold lower than the activity of the wild-type (WT) enzyme towards glutaryl-7-ACA. The report hinted at better results that were obtained using a different cephalosporin acylase, but these have not been reported yet [86]. Finally, screening for homologous genes in the microbial genome projects has revealed over 50 genes that are annotated as putative penicillin acylases. However, no experimental proof of the β-lactam acylase activity of a single one of these genes has been reported.

Chapter 1

Scope of the thesis

This thesis describes the search for an adipyl acylase, a β-lactam acylase that is capable of hydrolyzing adipyl-7-ADCA with high efficiency. The cephalosporin acylase from *Pseudomonas* SY-77 was used as the starting point. Random mutagenesis of the enzyme was performed by spiked oligo mutagenesis and error-prone PCR. Based on the obtained results, a stretch of three residues and two single residues were selected for saturation mutagenesis. The data from these random mutagenesis studies was combined with the 3D structure of the enzyme and the enzyme-substrate complex, enabling a rational saturation mutagenesis study of a small number of selected residues. Moreover, the best mutants of two positions were rationally combined. Finally, datamining of microbial genomes identified a cephalosporin acylase homolog in *P. aeruginosa* PAO1, which was shown to be an acyl-homoserine lactone acylase. Investigation of the physiological activity of this member of the acylase family suggests a role in the control of quorum sensing, a key process in bacterial infection.